

Protein domain screening system for structural genomics

○Eiko Seki¹, Takanori Kigawa^{1,2}, Natsuko Matsuda¹, Yoshihide Hayashizaki³, Shigeyuki Yokoyama^{1,2,4}

¹*Protein Research G., GSC, RIKEN Yokohama Institute, ²Cell Signal Lab., RIKEN Harima Institute at Spring-8, ³Genome Exploration Research G., GSC, RIKEN Yokohama Institute, ⁴Grad. Sch. of Sci., Univ. of Tokyo*

In RIKEN GSC Protein Research Group, we are heading for the systematical determination of the three-dimensional structures of protein folds. For this purpose, it is essential to prepare soluble and properly folded protein samples for structural analyses.

Here, we constructed a new experimental system, which enables us to select suitable samples for structural analyses. To begin with, target cDNA was randomly fragmented with Exonuclease III, to construct a deletion library of various lengths. For the 1st step screening, deleted fragments were expressed as C-terminal GFP-fusions in *E.coli*, and fluorescing clones were selected as candidates. For the 2nd step screening, those selected clones were expressed in cell free synthesis system, and screened on the basis of the intensity of GFP fluorescence.

We verified this screening system with mouse growth factor receptor-binding protein 2 (Grb2). The GFP fluorescence in the 2nd step screening was correlated with the solubility of deleted fragments. The terminals of soluble fragments coincided with the structural boundaries of Grb2. Thus, by using this system, we could screen protein samples for structural analyses, from the pool of various proteins. We are now applying this system to define structural domains of mouse cDNA.

序論

構造ゲノム科学は、タンパク質の立体構造解析を系統的かつ網羅的に行うものである。数万個といわれるタンパク質の立体構造は、数千種類の基本構造の組み合わせから形成されていると考えられており、我々はタンパク質の基本構造を網羅的に解明することにより、タンパク質の立体構造の全貌を明らかにしようとしている。この目的のためには、NMRやX線結晶構造解析を用いた立体構造解析に適した、構造をとり、かつ可溶性の高いタンパク質試料を、迅速に選び出す系の構築が重要である。

Waldo(1999)らの報告によると、タンパク質のC末端側にGFPを融合させた状態で大腸菌内で発現させることにより、その可溶性を検出することが可能である。しかしながら、大腸菌内で発現したタンパク質の蛍光強度を測定することは、煩雑であり、多検体同時処理に向いているとはいえない。そこで、我々は、cDNA deletion libraryのGFP-fusionと無細胞タンパク質合成系を組み合わせることにより、より正確かつ簡便に、可溶性のタンパク質ドメインを選び出す系の構築を行った。

タンパク質ドメイン選択法の開発

- 実験的タンパク質ドメイン選択法の必要性
 - 立体構造解析において、構造をとり、かつ可溶なタンパク質試料が必要である。
 - タンパク質の可溶性は、一次配列上からは予測出来ず、実験的手法でしか確認できない。
 - 構造ゲノム科学では多検体同時処理による構造決定が必要である。
- 必須条件
 - 簡便かつ迅速な方法
 - 立体構造をとったドメインの選択
 - 可溶性ドメインの選択
 - 大量合成系への簡便な移行

GFP融合体を利用した 可溶性サンプルのスクリーニング

GFP 融合体:

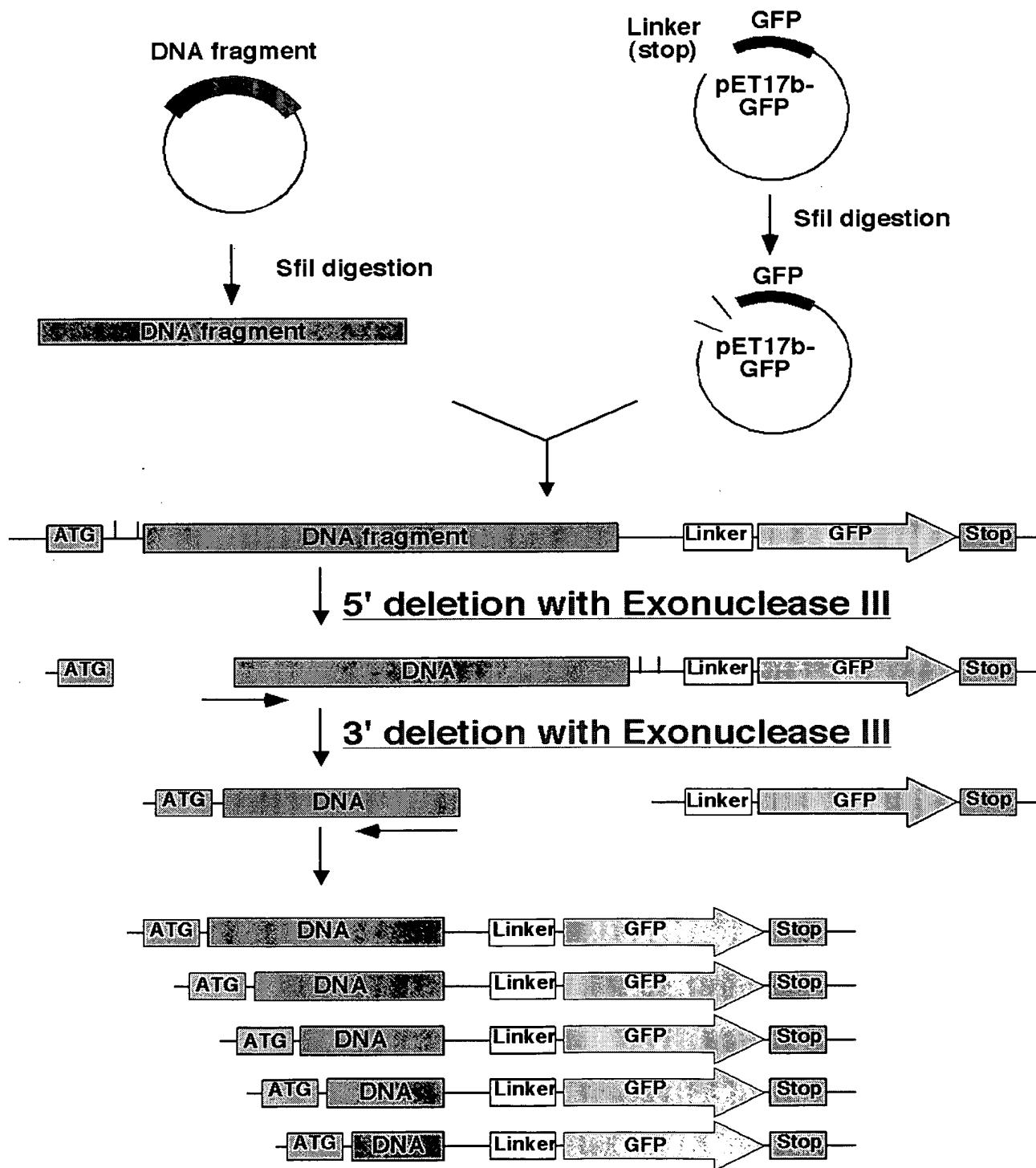
可溶性タンパク質をC末端側でGFP融合体にすると発光する
⇒ 難溶性タンパク質は発光しない

(Waldo G S et al Nat Biotechnol 17, 691-695 (1999))



様々な長さを含む均質で良質なdeletion libraryと組み合わせることで、効率よく構造解析に適したタンパク質ドメインをスクリーニングできる

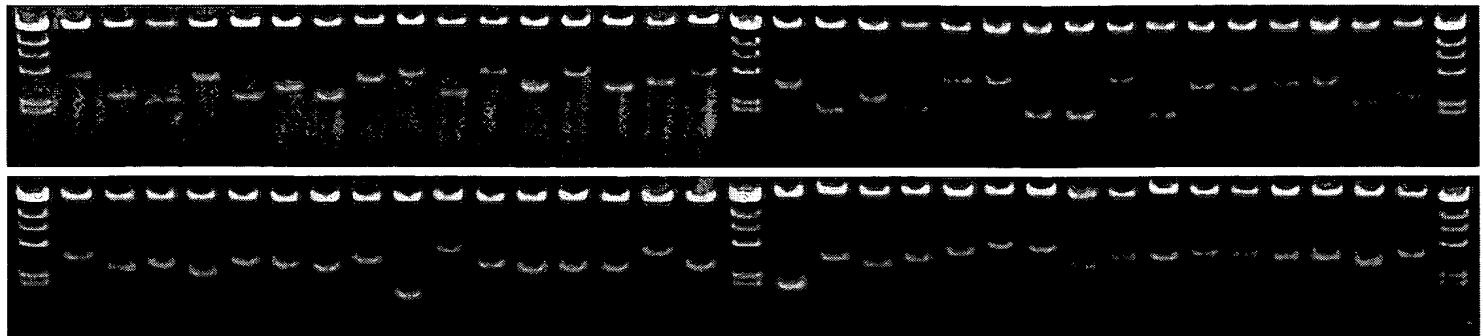
Exonuclease IIIを用いた Deletion Libraryの作成方法



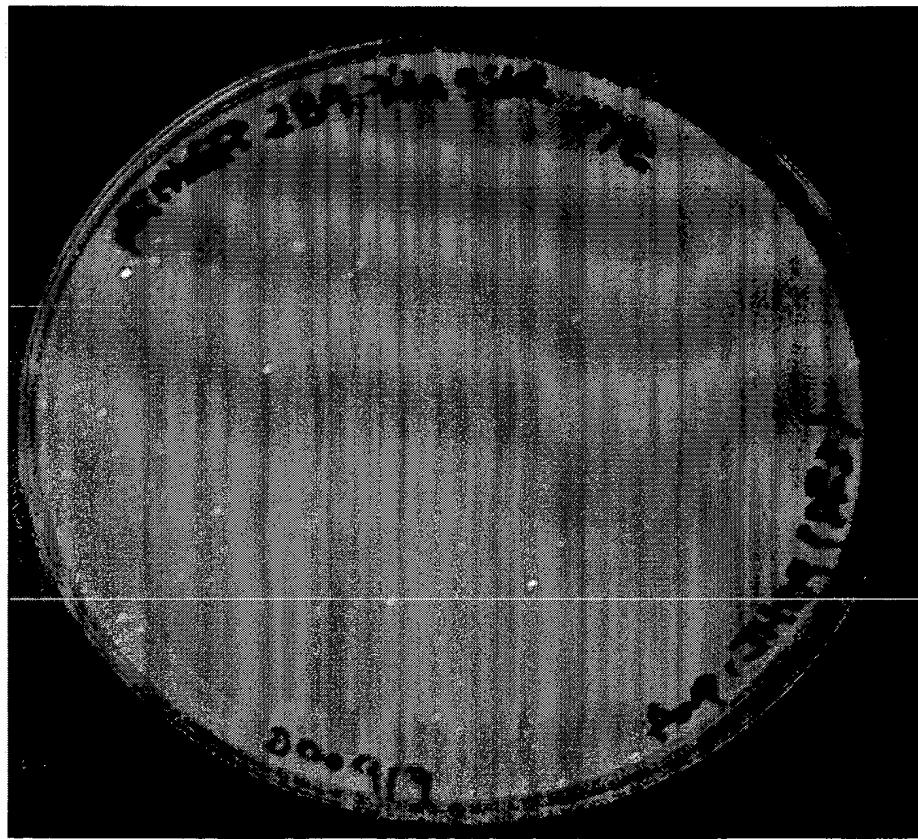
Target DNAの制限酵素サイト検索

Trial	OK		2Ken change				No enz.		O: 2sides deletion		△: 1 side deletion		X: no deletion							
	General	P.P	PCR成	R_SUP	Patl	EcoR	Small	KpnI	Spel	NheI	3'	3'NsI	5'Cla	5'Ks	EcoR					
X	14	IB2	76	ZX00001G13A	10740	3 498.9	411	##	##	##	##	##	##	##	##					
△	161	2F5	70	ZX00003C23A	16300	3 157.8	94.87	214	##	##	##	##	##	##	##					
X	2	1A2	66	ZX00001E07A	17121	3 360.8	245.8	238	##	##	##	##	##	##	##					
X	96	1H12	36	ZX00002H05A	17206	3 187.8	152.3	1331	587	1071	##	206	##	1638	##					
△	116	2B8	1	ZX0003M12A	20830	3 325.1	119.3	310	##	##	##	##	##	##	##					
X	70	1F10	38	ZX00002K10A	22088	3 534.1	98.9	258	##	##	##	##	##	##	600					
X	39	1D2	59	ZX00001H07	25649	3 424	303.1	1404	##	##	##	##	##	##	##					
O/△	144	2D12	64	ZX0003G10A	29655	3 440.2	129.5	53	##	##	##	##	##	##	##					
00201	O	56	1E8	48	ZX0002E09A	32885	3 215.6	202.8	##	##	##	##	##	##	##					
O/△	46	1D10	40	ZX0002D14A	33400	3 366.7	131.6	326	##	##	##	##	##	##	##					
O	187	2H7	23	ZX0003M04A	35674	3 288	216.5	703	##	452	##	##	##	##	##					
O/△	191	2H1	62	ZX0003H12A	38197	3 256.8	18.69	##	##	398	##	##	##	##	##					
O	123	2C3	20	ZX0002D10A	37969	3 193.1	153.3	615	350	##	##	##	##	##	##					
00201	O	184	2H4	72	ZX0003G09A	42083	3 57.8	50.12	##	25	1472	##	##	##	288					
00201	O	192	2H12	71	ZX0003P02A	43663	3 236.9	51.23	222	##	783	##	##	##	20					
O/△	119	2B11	42	ZX0003H12A	47702	3 134.5	87.5	##	##	##	138	##	##	##	703					
00201	O	73	1G1	21	ZX0001C13A	48119	3 390.1	320.1	635	##	##	659	##	##	##	403				
O	63	1G2	7	ZX0001H09A	50438	3 303.9	197.9	725	##	1221	##	##	##	##	##	##				
O	181	2H1	27	ZX0002P05A	52085	3 26.4	27.77	76	##	79	76	##	##	336	937	##				
O/△	143	2H11	77	ZX0003P02A	52599	3 152.4	16.01	1444	##	10	##	##	##	##	##	##				
O/△	75	1G3	32	ZX0001C08A	54264	3 440.2	384.4	140	##	612	89	##	##	951	579	##				
O/△	7	1A7	47	ZX0001F14A	60933	3 482.1	77.87	1391	##	170	##	##	##	##	##	##				
O	166	2F10	24	ZX0003L13A	61309	3 111.2	96.99	50	525	##	290	270	##	67	1421	##				
O/△	111	2B3	24	ZX0003D24A	64780	3 20	14.35	898	##	1085	##	56	563	##	170	1522	##			
00201	O	133	2D1	9	ZX0002J21A	68572	3 67.46	45.15	2165	2676	##	1829	##	##	##	1735	##			
O	74	1G4	4	ZX0001B03A	69242	3 385.5	190.73	1653	3158	##	##	120	##	##	##	583	##			
00201	O	112	2B4	14	ZX0003A21A	82257	3 158.1	123.2	6	##	##	##	1804	##	##	##	2562	##		
O/△	27	1C3	73	ZX0001M06A	20905	2 43.03	26.17	158	##	416	##	##	##	##	##	##	865	##		
X	67	1F7	44	ZX0001L04A	30333	2 268.3	171.9	838	##	##	##	324	##	##	##	##	##	##		
X	68	1F9	28	ZX0002B15A	31739	2 98.59	50.79	##	##	##	##	##	##	##	##	535	##	##		
O	53	1E5	53	ZX0001J05A	49239	2 235.2	67.51	397	##	##	##	##	##	##	##	997	##	##		
△	92	1H8	69	ZX0002E23A	49596	2 188.5	163.4	588	##	447	1326	##	##	##	##	1424	##	##		
O/△	21	1G9	20	ZX0002G11A	52939	2 15.19	139.9	229	##	21	53	##	##	##	##	681	##	##		
O/△	89	1H4	57	ZX0001B21A	71175	2 81.99	60.08	155	##	1052	78	##	##	##	##	1870	##	##		
O	1	1A1	56	ZX0001A05A	76132	2 107.3	62.98	56	##	1205	230	##	##	##	##	1698	##	##		
X	129	2C8	79	ZX0003F19A	13178	1 304.5	124.1	360	##	##	15	##	##	##	##	110	##	##		
X	183	2H3	55	ZX0002P10A	15051	1 280.9	68.85	##	##	##	##	##	##	##	##	535	##	##		
X	176	2G8	45	ZX0003F11A	24914	1 335.5	118.5	722	58	##	##	##	##	##	##	##	554	##	##	
O/△	137	2D5	39	ZX0003J15A	40009	1 216.5	136.3	509	##	928	##	##	##	##	##	##	661	##	##	
O/△	177	2G9	49	ZX0003J09A	45841	1 185.5	175.7	##	1056	336	##	##	##	##	##	515	##	687	##	
O	124	2C4	52	ZX0003C03A	46051	1 216.9	137.6	621	##	##	##	##	##	##	##	398	##	1117	##	
O	61	1F1	46	ZX0001C01A	51615	554.6	260.4	408	##	##	##	##	##	##	##	374	##	640	##	
O	140	2D8	25	ZX0003D20A	54874	3 390.4	188.1	##	817	##	##	##	##	##	##	529	##	##	##	
O	190	2H10	13	ZX0003N09A	61565	1 303.7	115.5	875	207	##	1535	##	##	##	##	1098	##	49	1201	##
O	190	2H10	13	ZX0003N09A	61565	1 303.7	115.5	875	207	##	1535	##	##	##	##	1098	##	49	1201	##
O	190	2H10	13	ZX0003N09A	61565	1 303.7	115.5	875	207	##	1535	##	##	##	##	1098	##	49	1201	##

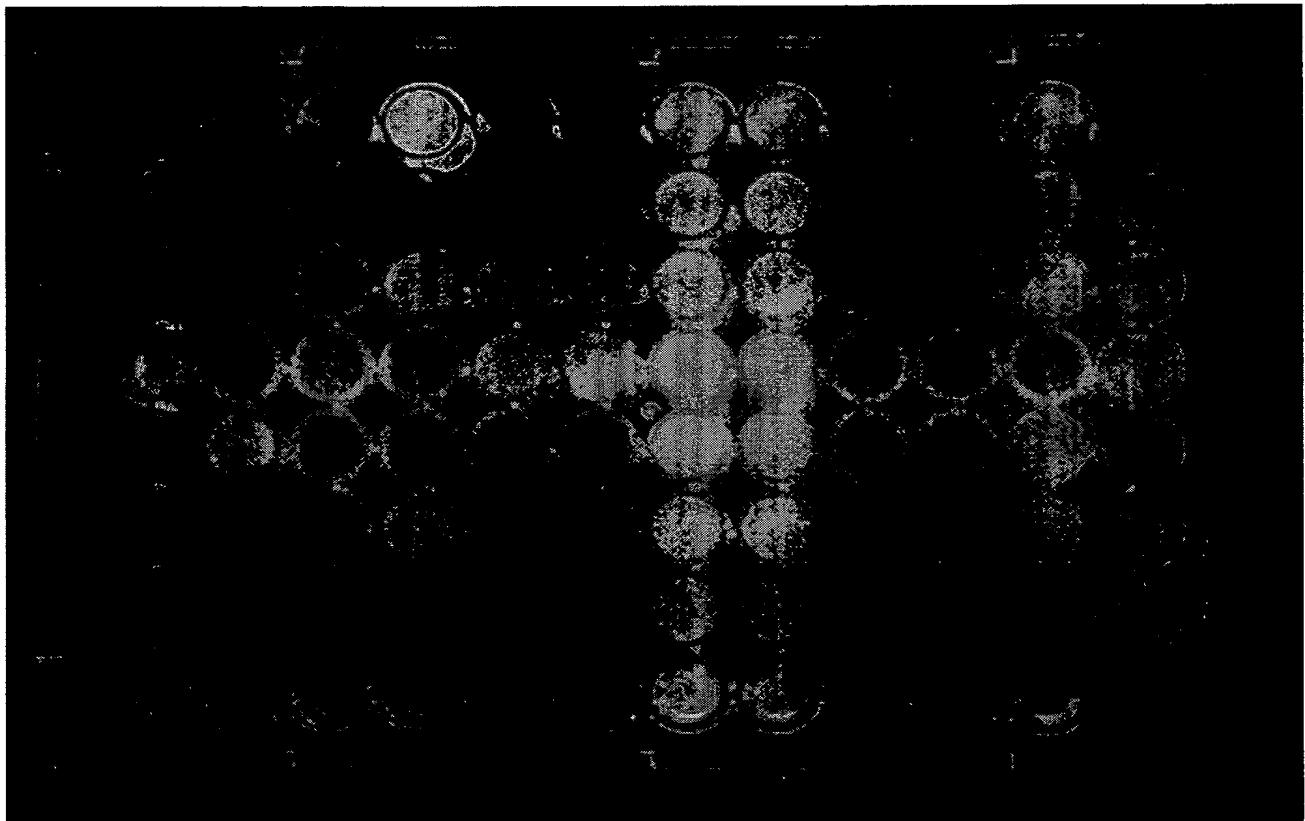
作成したDeletion Library



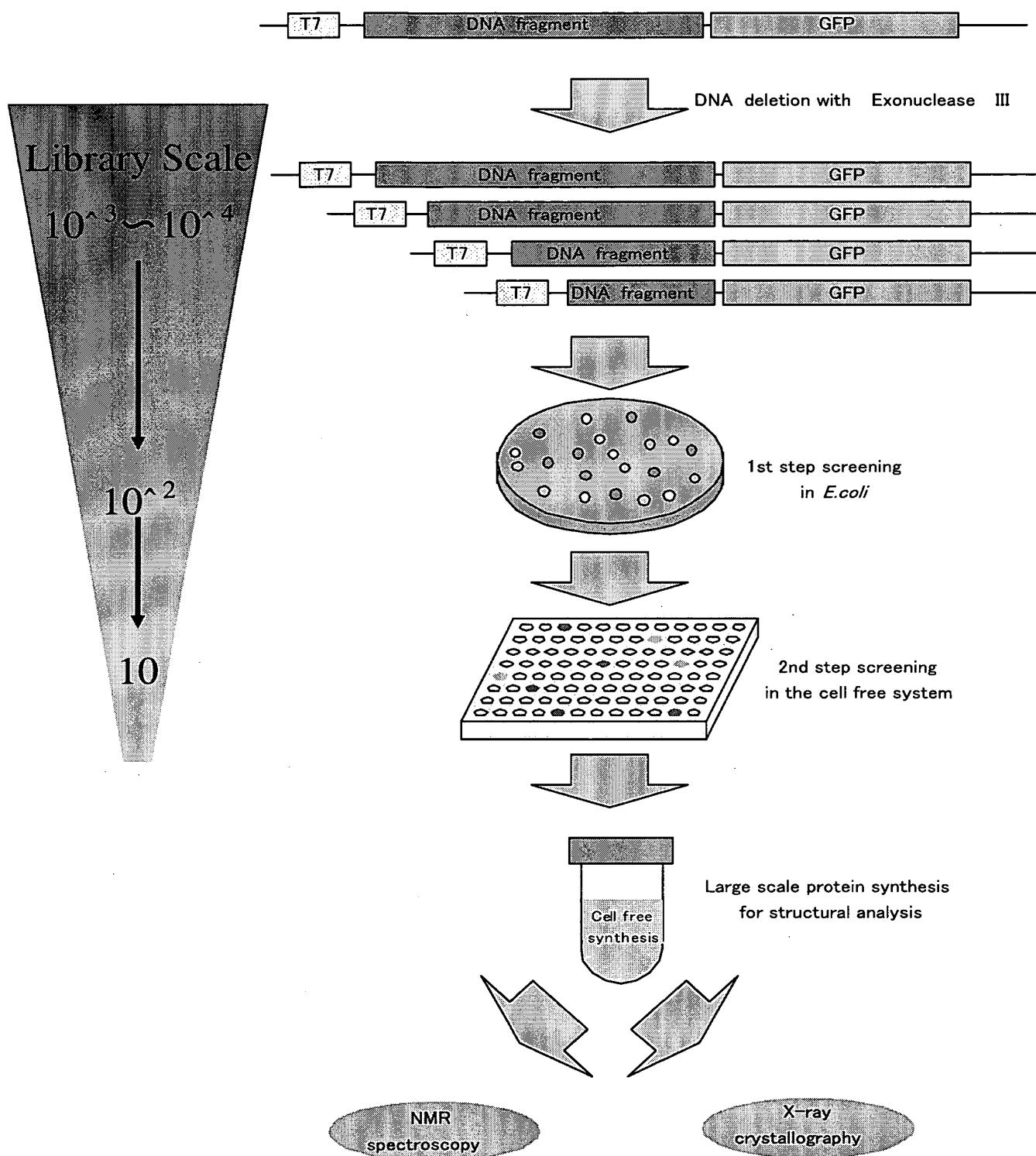
大腸菌での1段階目スクリーニング



無細胞合成での
2段階目スクリーニング



タンパク質ドメインの スクリーニングシステムの流れ



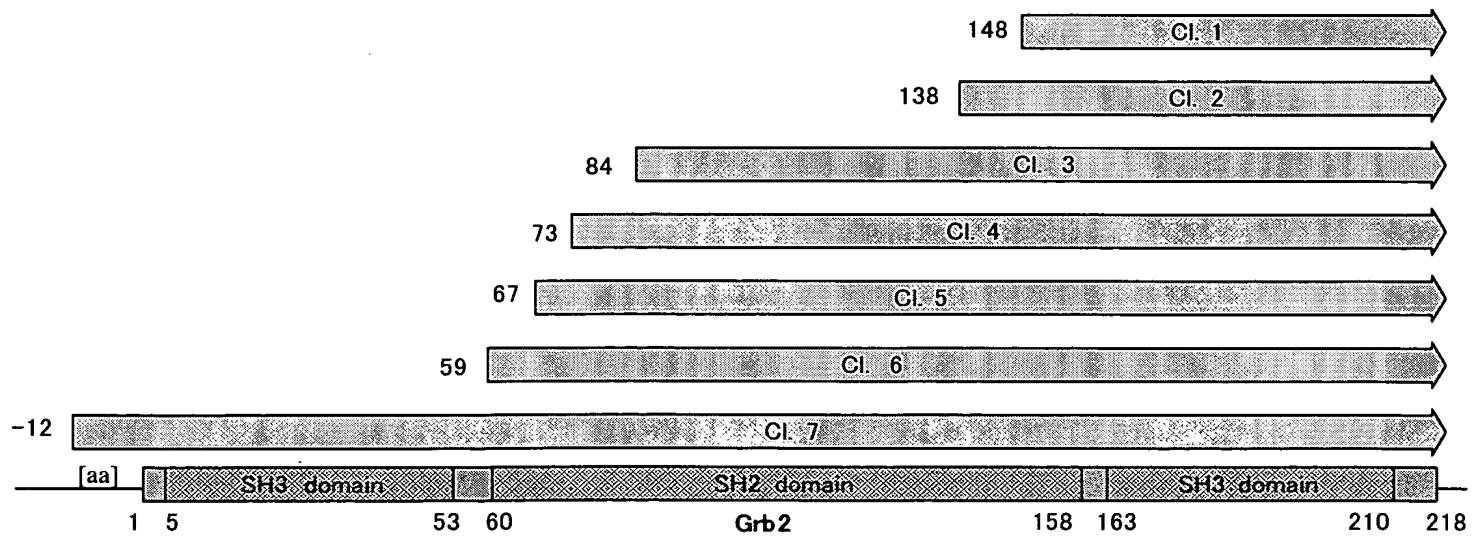
Grb2のドメインスクリーニング

- Grb2

- 全長で可溶性 GFP-fusionは発光
- SH3 SH2 SH3ドメインを含む
- 立体構造がすでに解析されている

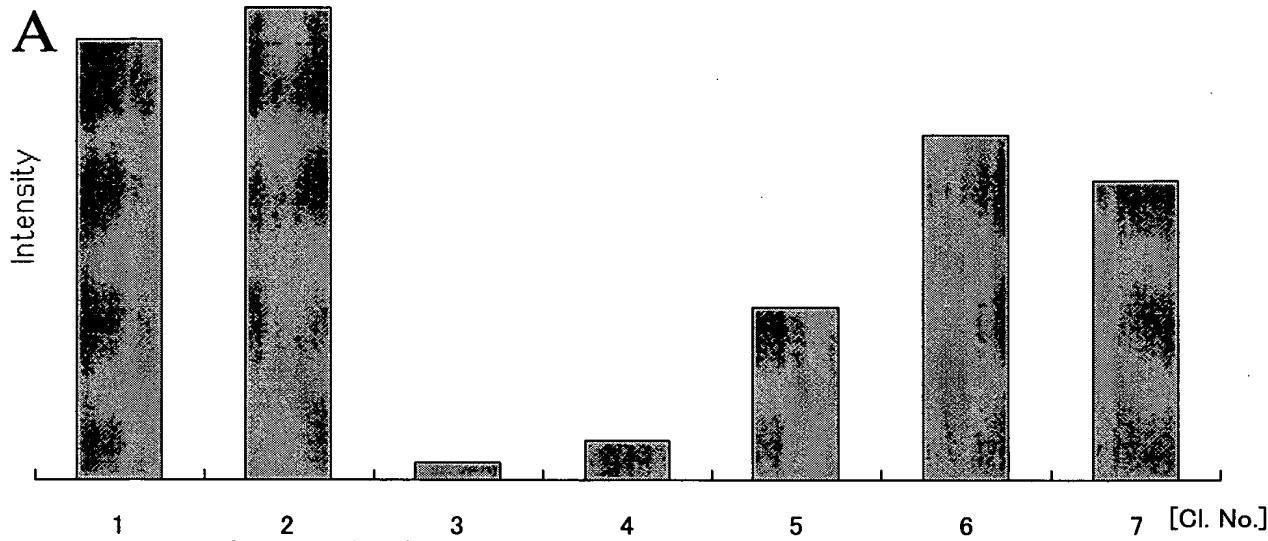
(Maignan, S. et. Al. Science 268, 291-293 1995)

スクリーニングにより選択された Deleted cloneの位置

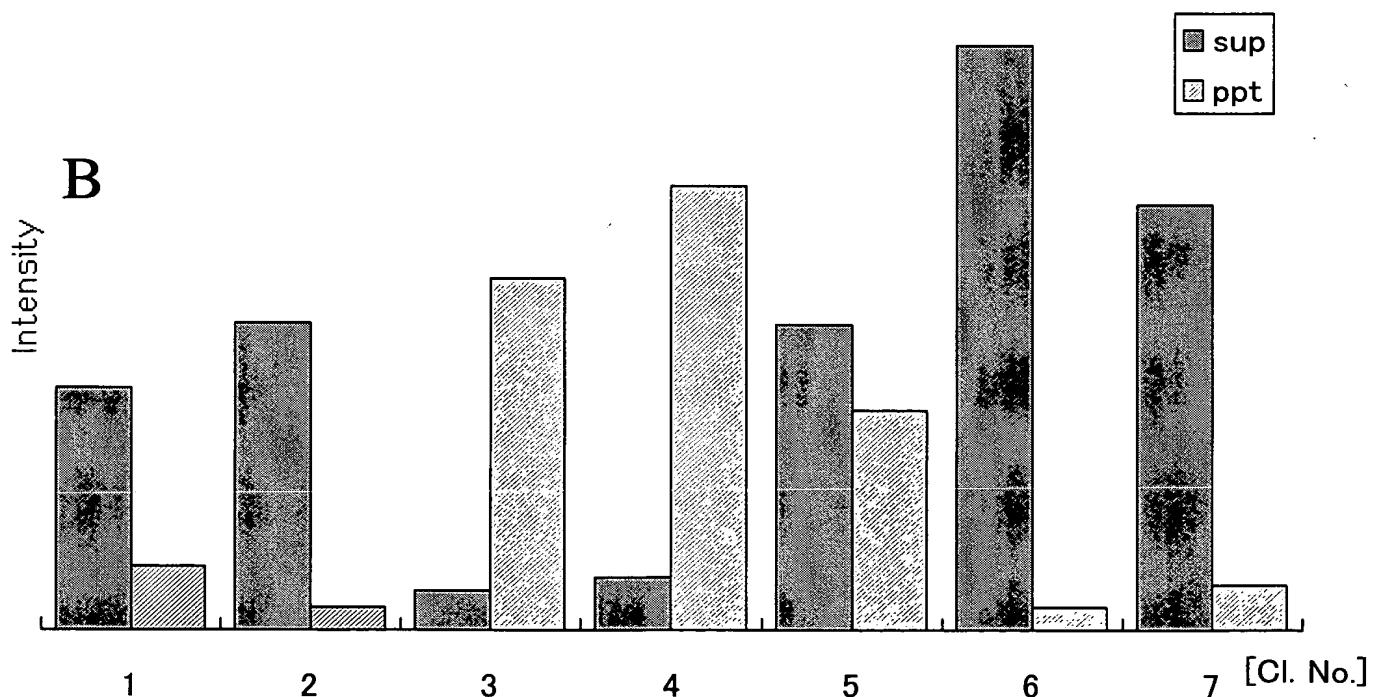


*ここでのドメインはアミノ酸配列の保存性から導き出された位置を示している

Grb2のDeleted cloneの解析(1)



A 無細胞合成による GFP-fusionの蛍光強度
(励起 485nm 発光 535nm)



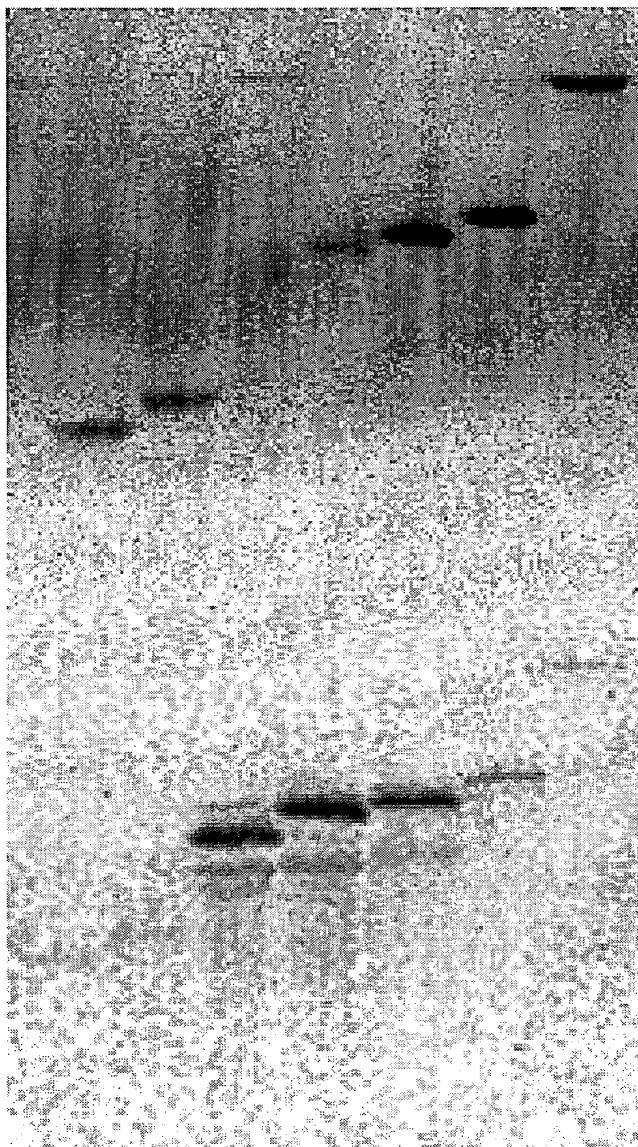
B 無細胞合成による単体での可溶性
(¹⁴C, scintillation counter)

Grb2のDeleted cloneの解析(2)

C

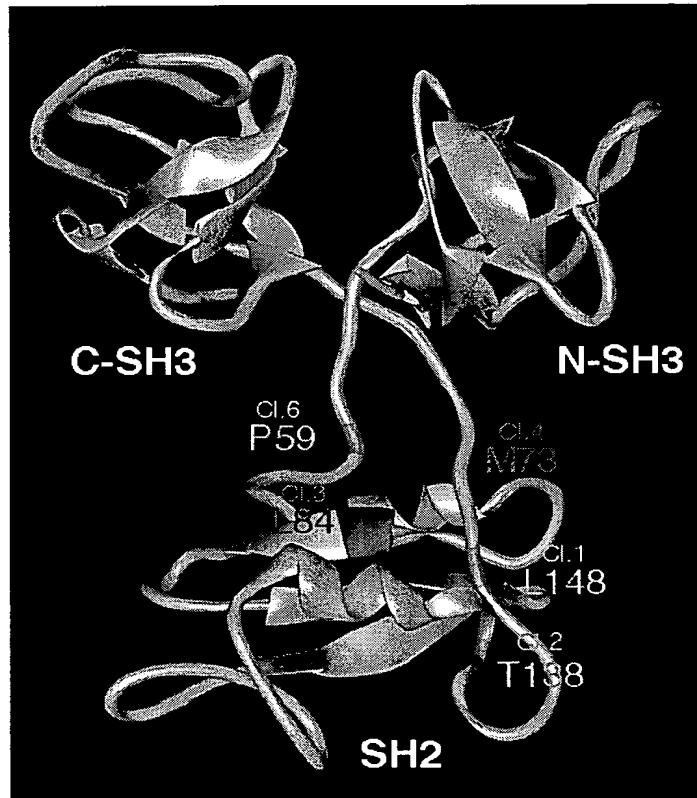
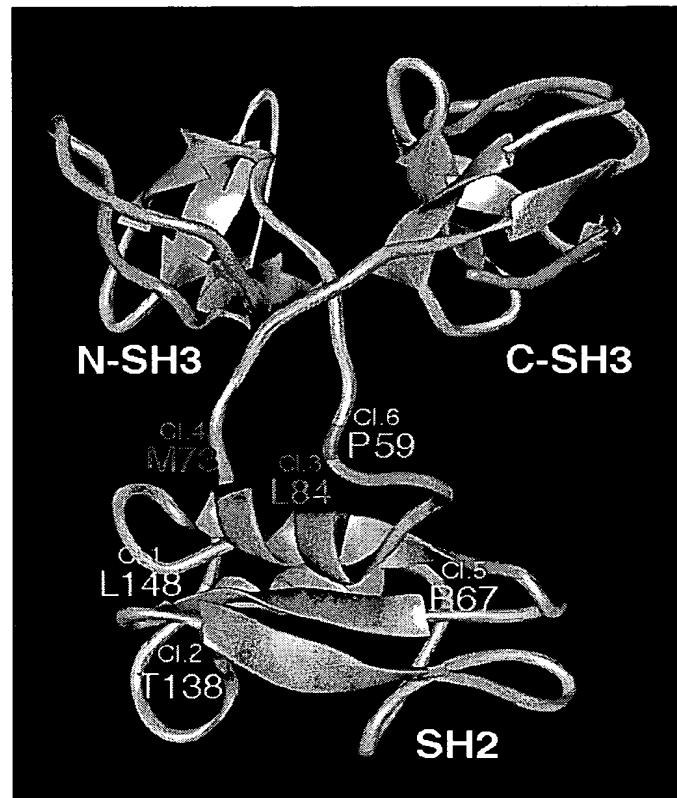
1 2 3 4 5 6 7

sup.



C. 無細胞合成による単体のSDS-PAGE
(¹⁴C,MacBas)

立体構造とDeletion部位の比較



■ : Soluble

■ : Partially soluble

■ : Insoluble

結論

- Exonuclease IIIを用いることで、均一でサンプリング範囲の広いDNA randomly deleted libraryを作成することが可能になった。
- DNA deleted libraryとタンパク質C末端側でのGFP-fusionを組み合わせ、さらに、大腸菌と無細胞系を用いて2段階でスクリーニングすることにより、可溶性の高いドメインを、効率よく選び出すことが可能になった。
- Grb2においては、本系が導き出した可溶性の高いfragmentの境界と、立体構造上のドメイン境界との間に強い相関があった。

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
 - FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.